



Calhoun: The NPS Institutional Archive

Theses and Dissertations

Thesis Collection

1982

A study of Foutz's multivariate goodness-of-fit test.

Linhart, Richard John

Monterey, California. Naval Postgraduate School

<http://hdl.handle.net/10945/20356>



Calhoun is a project of the Dudley Knox Library at NPS, furthering the precepts and goals of open government and government transparency. All information contained herein has been approved for release by the NPS Public Affairs Officer.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

<http://www.nps.edu/library>

LIBRARY
POSTGRADUATE SCHOOL
MONTEREY, CALIF. 93940

NAVAL POSTGRADUATE SCHOOL

Monterey, California



THESIS

A STUDY OF FOUTZ'S MULTIVARIATE
GOODNESS-OF-FIT TEST

by

Richard John Linhart, Jr.

March 1982

Co-Advisors:

T. Jayachandran
R. Franke
H. J. Larson

Approved for public release; distribution unlimited.

T204243
T204244

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Study of Foutz's Multivariate Goodness-of-Fit Test		5. TYPE OF REPORT & PERIOD COVERED Master's thesis; March 1982
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Richard John Linhart, Jr.		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE March 1982
		13. NUMBER OF PAGES 57
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Multi-variate goodness-of-fit test		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The empirical power of a new multivariate goodness-of-fit test proposed by Foutz (1980) is investigated. The test has been applied to Monte Carlo samples from bivariate and trivariate normal distributions with a variety of mean vectors and covariance matrices. The null hypothesis tested is that the sample is from a multi-variate normal distribution with <u>0</u> mean vector and covariance matrix the identity <u>I</u> . The		

(20. ABSTRACT Continued)

observed number of rejections in 5000 replications is used as the measure of effectiveness of the test. The results indicate that the Foutz test is quite capable of detecting mean and variance shifts but is not as powerful against covariance shifts.

Approved for public release; distribution unlimited.

A Study of Foutz's Multivariate
Goodness-of-Fit Test

by

Richard John Linhart, Jr.
Captain, United States Marine Corps
B.S., United States Naval Academy, 1972

Submitted in partial fulfillment of the
requirements for the degrees of

MASTER OF SCIENCE IN APPLIED MATHEMATICS
and
MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the
NAVAL POSTGRADUATE SCHOOL
March 1982

ABSTRACT

The empirical power of a new multivariate goodness-of-fit test proposed by Foutz (1980) is investigated. The test has been applied to Monte Carlo samples from bivariate and trivariate normal distributions with a variety of mean vectors and covariance matrices. The null hypothesis tested is that the sample is from a multivariate normal distribution with 0 mean vector and covariance matrix the identity I. The observed number of rejections in 5000 replications is used as the measure of effectiveness of the test. The results indicate that the Foutz test is quite capable of detecting mean and variance shifts but is not as powerful against covariance shifts.

TABLE OF CONTENTS

I.	INTRODUCTION -----	8
II.	THE FOUTZ TEST -----	10
III.	DESCRIPTION OF THE SIMULATION -----	18
IV.	RESULTS AND CONCLUSIONS -----	24
APPENDIX A:	USER REQUIREMENTS AND INPUT FORMAT FOR PROGRAM FOUTZ -----	46
APPENDIX B:	COMPUTER PROGRAM FOUTZ -----	48
	LIST OF REFERENCES -----	56
	INITIAL DISTRIBUTION LIST -----	57

LIST OF TABLES

I.	SAMPLE BIVARIATE DATA -----	12
II.	EMPIRICAL SIGNIFICANCE LEVEL OF THE FOUTZ F_n TEST -----	17
III.	APPROXIMATE CRITICAL VALUES FOR F_n TEST -----	17
IV.	NULL EMPIRICAL REJECTION LEVELS FOR THE BIVARIATE NORMAL DISTRIBUTION -----	27
V.	NULL EMPIRICAL REJECTION LEVELS FOR THE TRIVARIATE NORMAL DISTRIBUTION -----	28
VI.	REJECTION RATES FOR SHIFTS IN MEAN (BIVARIATE) -----	34
VII.	REJECTION RATES FOR SHIFTS IN MEAN (TRIVARIATE) -----	35
VIII.	REJECTION RATES FOR SHIFTS IN VARIANCE (BIVARIATE) -----	37
IX.	REJECTION RATES FOR SHIFTS IN VARIANCE (TRIVARIATE) -----	38
X.	REJECTION RATES FOR SHIFTS IN COVARIANCE (BIVARIATE) -----	40
XI.	REJECTION RATES FOR SHIFTS IN COVARIANCE (TRIVARIATE) -----	41
XII.	REJECTION RATES FOR MULTIPLE SHIFTS IN MEAN AND VARIANCE-COVARIANCE (BIVARIATE) -----	42
XIII.	REJECTION RATES FOR MULTIPLE SHIFTS IN MEAN AND VARIANCE-COVARIANCE (TRIVARIATE) ---	43
XIV.	REJECTION RATES FOR INCREASING SAMPLE SIZES (BIVARIATE) -----	44
XV.	REJECTION RATES FOR INCREASING SAMPLE SIZES (TRIVARIATE) -----	45

LIST OF FIGURES

1.	STATISTICALLY EQUIVALENT BLOCKS-- RECTANGULAR COORDINATES -----	13
2.	STATISTICALLY EQUIVALENT BLOCKS-- POLAR COORDINATES -----	14
3.	POWER CURVES FOR SHIFTS IN MEAN (BIVARIATE) ---	29
4.	POWER CURVES FOR SHIFTS IN MEAN (TRIVARIATE) --	30
5.	POWER CURVES FOR SHIFTS IN VARIANCE (BIVARIATE) -----	31
6.	POWER CURVES FOR SHIFTS IN VARIANCE (TRIVARIATE) -----	32
7.	POWER CURVES FOR SHIFTS IN COVARIANCE -----	33

I. INTRODUCTION

In statistical analysis, choosing the correct distribution to model available data is of importance. A class of procedures known as goodness-of-fit tests has been derived to test the hypothesis that a set of samples is from a given distribution. Many of these tests are readily available and are well known, such as the Chi-square or the Kolmogorov-Smirnoff (K.S.) goodness-of-fit test. These tests were designed for univariate distributions and are not usable as multivariate goodness-of-fit tests in their present form.

In 1980 Robert V. Foutz [Ref. 1] proposed a new multivariate goodness-of-fit test that will be called the Fn test in the sequel. In analogy to the K.S. test the Fn test compares a hypothesized cumulative distribution function (CDF) with a "continuous empirical distribution function" (CEDF) formed from sampled data. Foutz found the null distribution of the test to be distribution free as well as being independent of the number of variates p .

Foutz obtained an integral expression for the null distribution of the Fn test statistic, and closed form solutions for sample size 2 or 3 were provided. The complexity of the integral expression increases with sample size, and a normal approximation to the null distribution was given for use with larger sample sizes. Although

the Fn test was designed as a multivariate goodness-of-fit test it can also be used to fit univariate distributions. Franke and Jayachandran [Ref. 2] compared the empirical power of the Fn test with that for the Chi-square test and the K.S. test. The results indicated that the Fn test competes well with these other tests.

The power of the Fn test as a multivariate goodness-of-fit test is investigated in this thesis. A description of the Foutz test is given in Section II and the Monte Carlo methods of simulation are presented in Section III. The results and conclusions are in Section IV. A Fortran code for the application of the Fn test is available in the Appendix.

II. THE FOUTZ TEST

The Fn test for multivariate goodness-of-fit is based on a comparison of a hypothesized CDF with a continuous empirical distribution function (CEDF) derived from a sample. The first step in the determination of the CEDF is the construction of what are known as statistically equivalent blocks. A general method for determining statistically equivalent blocks, due to Anderson [Ref. 3], is described below.

Given a random sample $\underline{X}_1, \underline{X}_2, \dots, \underline{X}_{n-1}$ from a p-variate continuous distribution, select $n-1$ functions $h_k(\underline{X})$, $k = 1, 2, \dots, n-1$, not necessarily distinct, such that each $h_k(\underline{X})$ has a continuous distribution. These functions are referred to as cutting functions and will be used to partition the sample space into blocks. Let k_1, k_2, \dots, k_{n-1} be a permutation of $1, 2, \dots, n-1$. Order the \underline{X}_i 's according to $h_{k_1}(\underline{X})$ and define $\underline{X}(k_1)$ as the k_1 th order statistic. The sample space is partitioned into two blocks.

$$B_1 = \left\{ \underline{X}: h_{k_1}(\underline{X}) \leq h_{k_1}(\underline{X}(k_1)) \right\}$$

$$B_2 = \left\{ \underline{X}: h_{k_1}(\underline{X}) > h_{k_1}(\underline{X}(k_1)) \right\}.$$

At the second step if $0 < k_2 < k_1$ the $k-1$ \underline{X} 's in B_1 are ordered according to $h_{k_2}(\underline{X})$; $\underline{X}(k_2)$ is defined as the k_2 th in the ordering. Define a cut on B_1 obtaining 3 blocks as follows:

$$B_{11} = B_1 \cap \left\{ \underline{X}: h_{k_2}(\underline{X}) \leq h_{k_2}(\underline{X}(k_2)) \right\},$$

$$B_{12} = B_1 \cap \left\{ \underline{X}: h_{k_2}(\underline{X}) > h_{k_2}(\underline{X}(k_2)) \right\},$$

$$B_{20} = B_2.$$

Now consider the other alternative, $k_2 > k_1$. We rank the $((n-1)-k_1)$ \underline{X} 's in the second block B_2 according to $h_{k_2}(\underline{X})$ and let $\underline{X}(k_2)$ be the (k_2-k_1) th largest in the ranking. Defining a cut at $h_{k_2}(\underline{X}(k_2))$ we obtain the 3 blocks,

$$B_{10} = B_1,$$

$$B_{21} = B_2 \cap \left\{ \underline{X}: h_{k_2}(\underline{X}) \leq h_{k_2}(\underline{X}(k_2)) \right\},$$

$$B_{22} = B_2 \cap \left\{ \underline{X}: h_{k_2}(\underline{X}) > h_{k_2}(\underline{X}(k_2)) \right\}.$$

The process is continued until all the cutting functions are exhausted. This results in a partition of the sample space into n statistically equivalent blocks, which are denoted by B_i , $i = 1, \dots, n$.

In the univariate case an intuitively appealing choice for the cutting functions is the identity function viz., $h(X) = X$ for all k . The resulting statistically equivalent blocks are then $(-\infty, X(1)], (X(1), X(2)], \dots, (X(n-1), +\infty)$ where $X(j)$ is the j th order statistic. The multivariate analogue is to choose

individual coordinates as cutting functions, viz., $h_k(\underline{X}) = \underline{X}^{(j)}$, the j th coordinate of \underline{X} . An example illustrating the construction of the blocks in the bivariate case is given below for a sample of size 8.

Let $(2,4,6,8,1,3,5,7)$ be the permutation vector K . Define $h_k(\underline{X}) = \underline{X}^{(1)}$, the first coordinate of \underline{X} , for $k = 2,4,6,8$ and $h_k(\underline{X}) = \underline{X}^{(2)}$, the second coordinate, for $k = 1,3,5,7$. Figure 1 gives a graphical representation of the rectangular coordinate method of forming blocks and Figure 2 is the representation for the polar coordinate method. The random sample that was used in both figures is found in Table I.

TABLE I: SAMPLE BIVARIATE DATA

N = 8								
Observation Coordinate	1	2	3	4	5	6	7	8
1	-3.54	2.25	-1.00	.71	2.00	-.75	-2.25	0.00
2	0.00	-2.25	0.50	.00	1.25	-1.50	-1.50	-0.50

The first element of the permutation vector is $k = 2$ and $h_2(\underline{X}) = \underline{X}^{(1)}$, therefore $x_2^{(1)}$ is defined to be the second smallest first coordinate. This partitions the sample space into two blocks,

$$B_1 = \left\{ X: X^{(1)} \leq x_2^{(1)} \right\},$$

$$B_2 = \left\{ X: X^{(1)} > x_2^{(1)} \right\}.$$

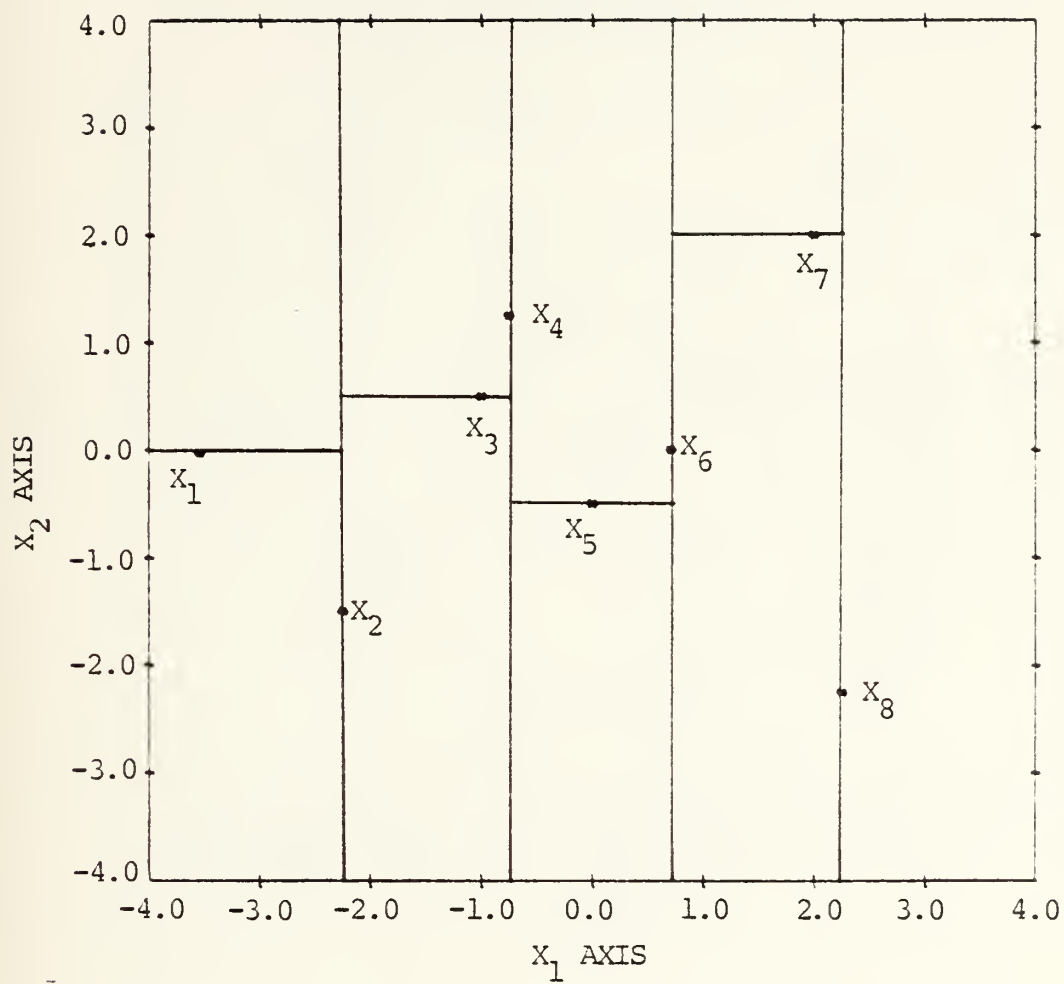


FIGURE 1: STATISTICALLY EQUIVALENT BLOCKS--
RECTANGULAR COORDINATES

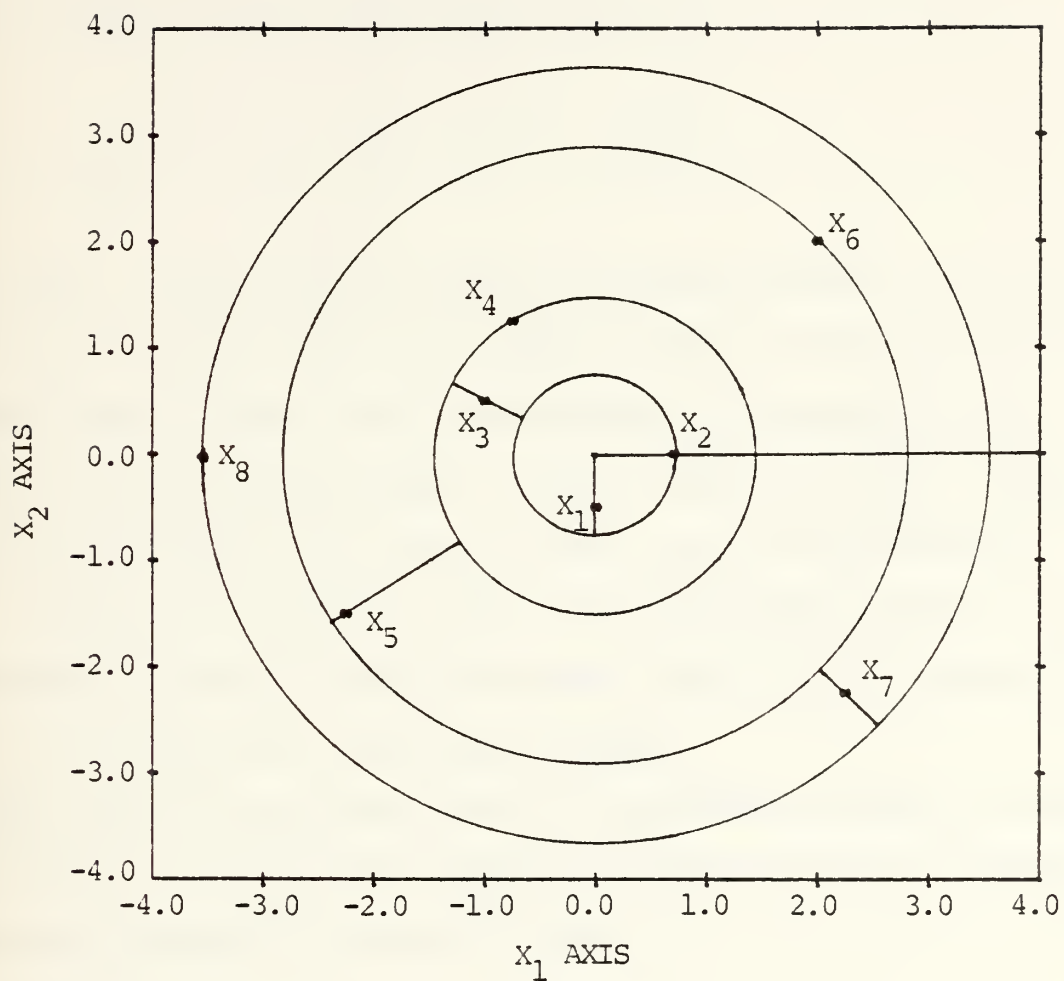


FIGURE 2: STATISTICALLY EQUIVALENT BLOCKS--
POLAR COORDINATES

The second element of the permutation vector is $k_2 = 4$, $h_4(\underline{X}) = \underline{X}^{(1)}$ and $k_2 > k_1$. Hence the block B_2 is partitioned into two sub-blocks,

$$B_{21} = B_2 \cap \left\{ X: X^{(1)} \leq X_2^{(1)} \right\},$$

$$B_{22} = B_2 \cap \left\{ X: X^{(1)} > X_2^{(1)} \right\},$$

where $X_2^{(1)}$ is the second largest coordinate among the \underline{X} 's in block B_2 . At this stage the sample space is partitioned into three blocks. Next, the third element of the permutation vector and the corresponding cutting function define another partition of one of the three blocks into two sub-blocks. This process is continued until the permutation vector is exhausted, at which stage the sample space will be partitioned into 9 statistically equivalent blocks.

The CEDF is now constructed by spreading a mass $1/n$ within each block. If H_0 is the hypothesized CDF and H_n the CEDF, the test statistic F_n takes the form

$$F_n = \sup_{\underline{X}} |H_n(\underline{X}) - H_0(\underline{X})|. \quad (1)$$

Let D_i , $i = 1, 2, \dots, n$, be the probability contents of the blocks B_i under the null hypothesis H_0 , i.e., $D_i = \int_{B_i} dH_0(\underline{x})$.

A computational form of the Foutz test statistic is,

$$F_n = \sum_{i=1}^n \text{Max} \left(0, \frac{1}{n} - D_i \right). \quad (2)$$

Foutz gave the following representation for the cumulative distribution of the test statistic

$$P(F_n < x) = \int_{-\infty}^x \dots \int_{-\infty}^x g_n(\delta_1, \delta_2, \dots, \delta_{n-1}) d\delta_1 d\delta_2, \dots, d\delta_{n-1}; \quad (3)$$

where

$$g_n(\delta_1, \delta_2, \dots, \delta_{n-1}) = n!(n-1)!$$

for

$$\frac{1}{n} \geq \delta_1 > (\delta_2 - \delta_1) > \dots > (\delta_{n-1} - \delta_{n-2}) > -\delta_{n-1}.$$

The evaluation of this integral is cumbersome and has not been carried out for $n > 5$. Foutz has therefore derived a large sample normal approximation given by

$$\lim_{n \rightarrow \infty} P[F_n \leq x] = \Phi \left[\frac{n^{(1/2)} (x - e^{-1})}{(2e^{-1} - 5e^{-2})^{1/2}} \right]. \quad (4)$$

To check the accuracy of the normal approximation, Franke and Jayachandran [Ref. 4] generated 80,000 samples of sizes 20, 30 and 50. Table II contains the empirical significance

TABLE II: EMPIRICAL SIGNIFICANCE LEVEL OF THE FOUTZ F_n TEST

Sample Size	20	30	50
Normal Significance Level			
.10	.0757	.0800	.0859
.05	.0372	.0399	.0428
.01	.0082	.0083	.0093

levels, when the normal approximation was used to determine the critical values for the F_n test.

It is clear that the rejection rates given in Table II are consistently lower than the nominal values. More accurate critical values were therefore determined from the 80,000 F_n values and are presented in Table III.

TABLE III: APPROXIMATE CRITICAL VALUES FOR F_n TEST

Sample Size	20	30	50
Significance Level			
.10	.42714 (.43586)	.41903 (.42383)	.40816 (.41150)
.05	.44865 (.45513)	.43553 (.43969)	.42116 (.42386)
.01	.48659 (.49127)	.46579 (.46944)	.44487 (.44706)

Values in parentheses are those obtained from the normal approximation given by Foutz.

III. DESCRIPTION OF THE SIMULATION

In order to check the efficacy of the Foutz test as a multivariate goodness-of-fit test a simulation was run to generate sample data from various bivariate and trivariate normal distributions. The hypothesis tested in each case is that the sample is from a multivariate normal distribution with mean vector $\underline{0}$ and covariance matrix the identity \underline{I} . Rectangular and the polar/spherical method of blocking were both used and compared as to their effect in each case.

To validate the blocking schemes, the null hypothesis is tested against data generated from the distribution $N(\underline{0}, \underline{I})$. Bivariate and trivariate sample sizes of 20, 30 and 50 are used to compute the F_n statistic which is then compared to the empirical critical levels found in Table III. Rejection rates are based on the number of rejections in 20,000 replications for each sample size. Comparing the null rejection rates to the nominal significance level used, as shown in Table III, provides evidence supporting both blocking methods as all null rejection rates are close to the significance level used.

The empirical power of the test was then investigated by varying the distribution tested. This investigation is accomplished in three different ways. First, the mean is

shifted away from the $\underline{0}$ vector while leaving the covariance as the identity matrix. This is done to investigate the ability of the test to detect location shifts. The covariance matrix is then changed from the identity while leaving the mean as the $\underline{0}$ vector. This is accomplished by changing the diagonal elements alone to investigate variance shifts and then shifting the off diagonal elements by themselves to check the effect of covariance shifts. A primary sample size of 20 was chosen for comparison and 5000 replications were used to compute rejection rates for each distribution tested. Mixing of the three types of shifts is also simulated to investigate the possible confounding effects of the three shifts. Finally sample sizes of 30 and 50 are run on a few of the distributions to determine the effect of increasing the sample size.

The various multivariate normal distributions are simulated in the following manner. Univariate normal(0,1) pseudorandom deviates are obtained from the LLRANDU series by Lewis [Ref. 5] and grouped to form a multivariate $N(\underline{0}, \underline{I})$ p-variate vector. Taking the \underline{X}^* so formed, the p-variate $N(\underline{0}, \underline{I})$ vector random variable is transformed by

$$\underline{C}^{-1} \underline{X}^* + \underline{\mu} = \underline{X} , \quad (1)$$

where

$$\underline{C}' \underline{\Sigma} \underline{C} = \underline{I} ,$$

resulting in an \underline{X} which is distributed as $N(\underline{\mu}, \underline{\Sigma})$. The Foutz test is then applied to each of the samples consisting of $(n-1) \underline{X}$ s.

An example using a bivariate sample helps illustrate the blocking procedure used. Let $\underline{X}_1, \underline{X}_2, \dots, \underline{X}_{n-1}$, be the simulated bivariate sample. The first cut is made on $\underline{X}_1^{(1)}$ or the first coordinate of the first vector \underline{X}_1 . Two blocks are formed,

	First Coordinate	Second Coordinate
B_1	$(-\infty, \underline{X}_1^{(1)}]$	$(-\infty, +\infty)$
B_2	$(\underline{X}_1^{(1)}, +\infty)$	$(-\infty, +\infty)$

\underline{X}_2 is taken next and determined to be contained in block B_1 or B_2 . Suppose \underline{X}_2 is in block B_2 . B_2 is then partitioned by $\underline{X}_2^{(2)}$ or the second coordinate of sample \underline{X}_2 . Three blocks are now defined as,

	First Coordinate	Second Coordinate
B_{10}	$(-\infty, \underline{X}_1^{(1)}]$	$(-\infty, +\infty)$
B_{21}	$(\underline{X}_1^{(1)}, +\infty)$	$(-\infty, \underline{X}_2^{(2)}]$
B_{22}	$(\underline{X}_1^{(1)}, +\infty)$	$(\underline{X}_2^{(2)}, +\infty)$

This procedure is continued by examining the next vector in the random sample, locating the block that it is contained in and partitioning the block by the designated coordinate. The coordinate cutting functions used are alternated starting with the first coordinate for the first cut. Coordinate ranges, as shown, are used to designate blocks and the process is continued until n blocks are so defined. Given any random sample this method can be shown to be equivalent to a unique permutation vector K and a set of cutting functions $\{h_k\}$ as defined in Section II.

After the formation of the statistically equivalent blocks, each block has the probability content of $1/n$ and must be compared to the hypothesized content using the statistic

$$F_n = \sum_{i=1}^n \max[0, \frac{1}{n} - D_i]. \quad (2)$$

D_i , the probability content of each block, under the null hypothesis, is defined by the integral of the null density over the block. The integral of the multivariate normal $(\underline{0}, \underline{I})$ over a rectangular block yields

$$D_i = \int \dots \int_{B_i} (2\pi)^{\frac{-p}{2}} e^{-(1/2) \underline{x}' \underline{I} \underline{x}} d\underline{x}. \quad (3)$$

This reduces to the product of the marginal densities which may be easily evaluated with many available routines, eliminating the need for numerical integration.

In spherical coordinates D_i is represented by

$$D_i = \int_{\phi_1}^{\phi_2} \int_{\theta_1}^{\theta_2} \int_{\rho_1}^{\rho_2} 2\pi^{(-3/2)} e^{(-1/2)\rho^2} \sin(\phi)\rho^2 d\rho d\theta d\phi. \quad (4)$$

Upon separation,

$$D_i = \int_{\phi_1}^{\phi_2} (1/2) \sin \phi d\phi \int_{\theta_1}^{\theta_2} (2\pi)^{-1} d\theta \int_{\rho_1}^{\rho_2} \frac{2\rho^2 e^{(-1/2)\rho^2} d\rho}{(2\pi)^{1/2}}. \quad (5)$$

Noting that with a change of variables the third integrand is a Chi-square density with 3 degrees of freedom, we may use a closed form expression to evaluate D as follows:

$$D_i = \left[\frac{1}{2} (\cos \phi_2 - \cos \phi_1) \right] \times \left[\frac{1}{2\pi} (\theta_2 - \theta_1) \right] \times [\chi_{3df}^2(\rho_2) - \chi_{3df}^2(\rho_1)] \quad (6)$$

where

$$\chi_{3df}^2(\rho_i) = P[\chi_{3df}^2 \leq \rho_i], \quad i = 1, 2.$$

For bivariate data the use of polar coordinates leads to similar simplification leaving D_i in the form

$$D_i = \frac{1}{2\pi}(\theta_2 - \theta_1) \times [\chi_{2df}^2(R_2)] - [\chi_{2df}^2(R_1)]. \quad (7)$$

After the calculation of the probability contents D_i for the n blocks, equation (2) is used to evaluate the F_n statistic for each generated sample. The statistic is then compared to the critical values found in Table III to decide if the null hypothesis is accepted or rejected. Rejection rates are defined by the number of rejections divided by the number of replications in a given run. The rejection rates thereby define an empirical power for the simulated distribution.

The major component of the Fortran simulation program used to evaluate the Foutz statistic for a given sample is available in the Appendix. It has been adapted for use for sample sizes up to 50, with redimensioning being needed for larger sample sizes. The program is applicable for fitting data from any hypothesized multivariate normal distribution and provides the F_n statistic as computed by both blocking methods presented. The code is self-contained except for three IMSL routines, LUDECP, MDNOR, and MDCH [Ref. 6]. These subroutines provide matrix decomposition, univariate normal probabilities and chi-square probabilities, respectively, and must be available or substituted prior to utilization of the program.

IV. RESULTS AND CONCLUSIONS

The results of the simulation are summarized in Tables IV-XIV. Rejection rates are given by the distribution tested and the significance level used. Empirical power curves are presented in Figures 3-8. Rejection rates are plotted against the magnitude of the shift in mean, variance and covariance for the distribution tested. All power curves are based on 5000 replicated samples and were compared at the $\alpha = .05$ significance level.

The results for the case in which the distribution of the samples is the same as the hypothesized distribution viz., $N(\underline{0}, \underline{I})$ are given in Tables IV and V. The rejection levels obtained are close to the nominal significance level for both blocking methods. No distinct pattern of variation about the prescribed levels is discernible for either method, as expected.

The rejection rates for mean shifts are given in Tables VI-VII and Figures 3-4. Shifts in the mean vector are detected well; a shift of one standard deviation in a single coordinate resulted in a 60% rejection rate for bivariate or trivariate data. Greater shifts in mean led to even higher rejection rates. The rectangular method of blocking consistently gave about a 10% improvement over the polar/spherical method in detecting mean shifts.

Results for variance shifts are contained in Tables VIII and IX and the power curves are given in Figures 5 and 6. The Foutz test did not detect small variance shifts very well but the performance of the test was far better for larger shifts or shifts in more than one coordinate. No one method of blocking performed better in all cases but in general the polar/spherical method seemed to outperform the rectangular method for detecting variance shifts.

The results for changes in covariance are summarized in Tables X, XI and Figure 7. Covariance shifts are not detected well for either blocking method except for highly correlated data with the correlation coefficient equal to .9. The polar/spherical coordinate blocking method appeared to perform a little better than the rectangular coordinate method of blocking, but in general the simulation revealed that the Fn test is not very powerful against covariance shifts.

The empirical power for combinations of shifts in mean and variance or covariance are presented in Tables XII and XIII. Entries are based on an $\alpha = .05$ significance level and are tabled by the mean vector and covariance matrix of the sample data. Entries farther down and to the right correspond to greater shifts in mean and variance/covariance and are generally larger, as is to be expected. There are no apparent confounding problems due to shifts in both

parameters. The rectangular method of blocking, however, did outperform the polar/spherical method for most cases of multiple shifts.

The results indicative of the effect of increasing the sample size are summarized in Tables XIV and XV. Results for sample sizes of 20, 30, and 50 are given for some representative cases. The tables reveal higher rejection rates for larger sample sizes with increases being comparable for both blocking methods.

This study was limited to the two and three variate normal distribution. There are many problems for further research. Of primary concern is the generation of percentage points of F_n for various values of n . The intractability of the problem of obtaining the exact distribution requires an empirical approach to finding a correction to the asymptotic approximation given by Foutz. Since the use of coordinates as cutting functions worked well, the method should be tried for other distributions and higher dimensions.

In conclusion, the F_n test is found to be a viable option for testing goodness-of-fit of multivariate normal distributions. These encouraging empirical results indicate further study should be conducted to explore the potential of this test for other distributions.

TABLE IV: NULL EMPIRICAL REJECTION LEVELS FOR
THE BIVARIATE NORMAL DISTRIBUTION

Significance Level	.01	.05	.10
Blocking Method			
		N = 20	
Rectangular	.0098	.0488	.0940
Polar	.0096	.0482	.1020
		N = 30	
Rectangular	.0110	.0510	.0944
Polar	.0082	.0454	.0890
		N = 50	
Rectangular	.0120	.0498	.0950
Polar	.0098	.0484	.0958

BASED ON 20,000 REPLICATIONS

TABLE V: NULL EMPIRICAL REJECTION LEVELS FOR
THE TRIVARIATE NORMAL DISTRIBUTION

Significance Level	.01	.05	.10
Blocking Method			
		N = 20	
Rectangular	.0104	.0440	.0982
Spherical	.0120	.0518	.1048
		N = 30	
Rectangular	.0114	.0480	.0956
Spherical	.0140	.0484	.0914
		N = 50	
Rectangular	.0098	.0484	.0960
Spherical	.0088	.0478	.0914

BASED ON 20,000 REPLICATIONS

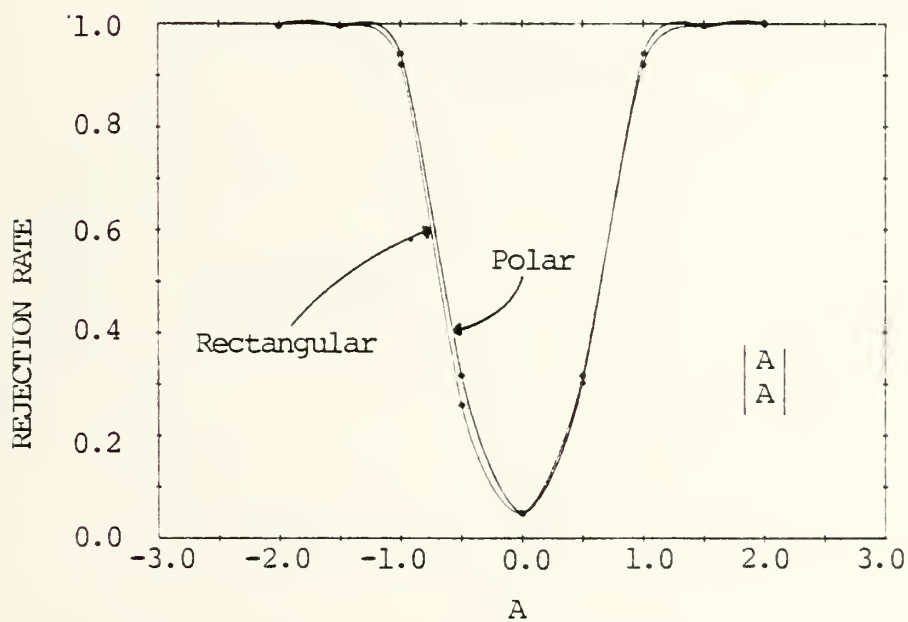
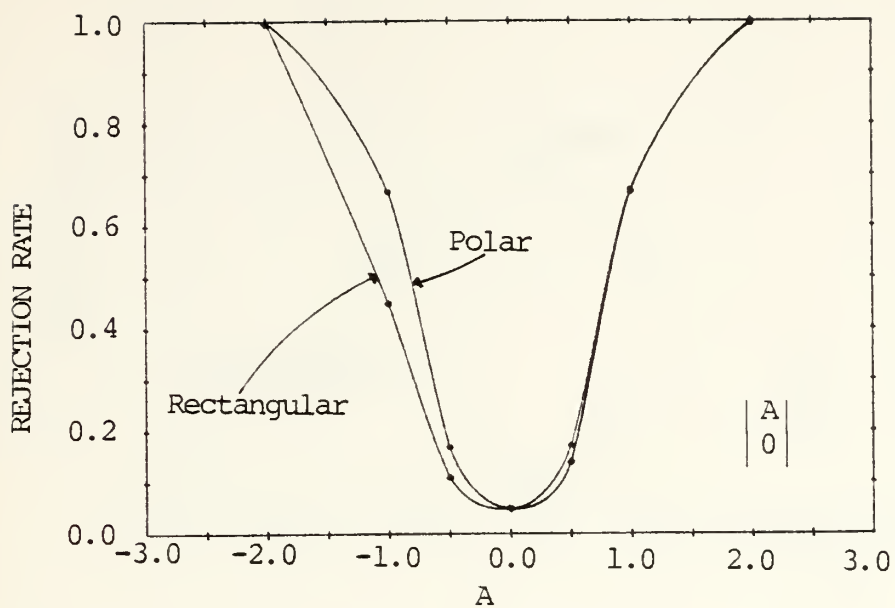


FIGURE 3: POWER CURVES FOR SHIFTS IN MEAN
(BIVARIATE)

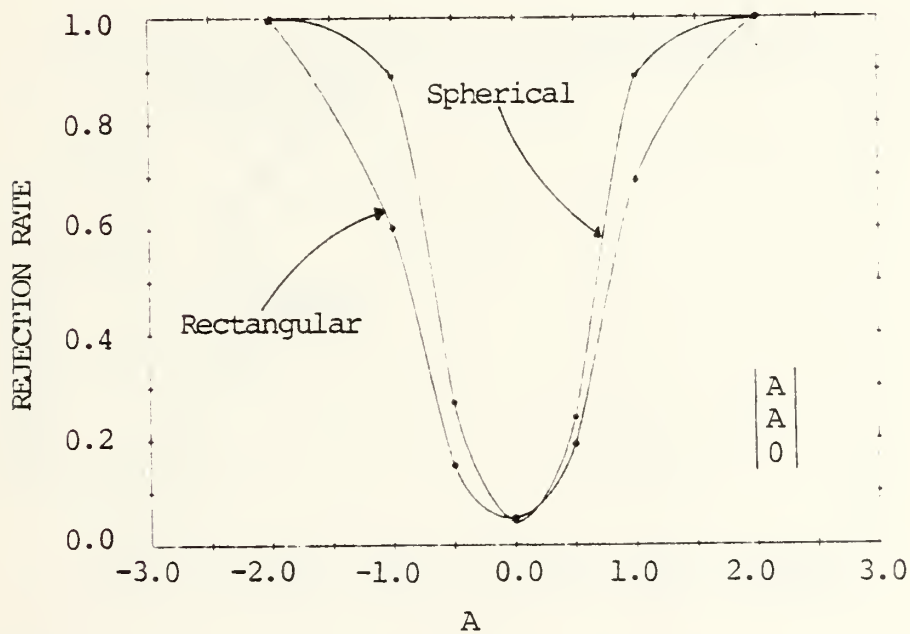
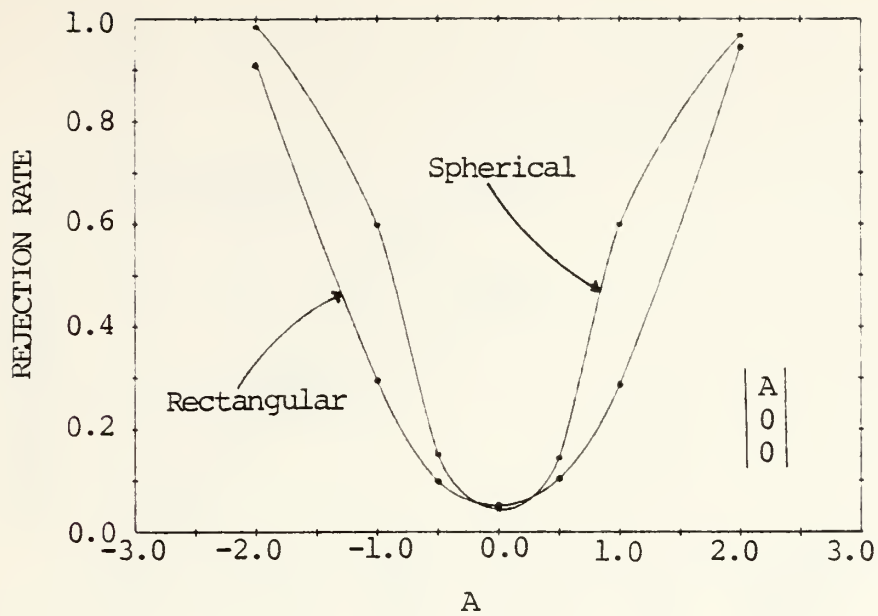


FIGURE 4: POWER CURVES FOR SHIFTS IN MEAN
(TRIVARIATE)

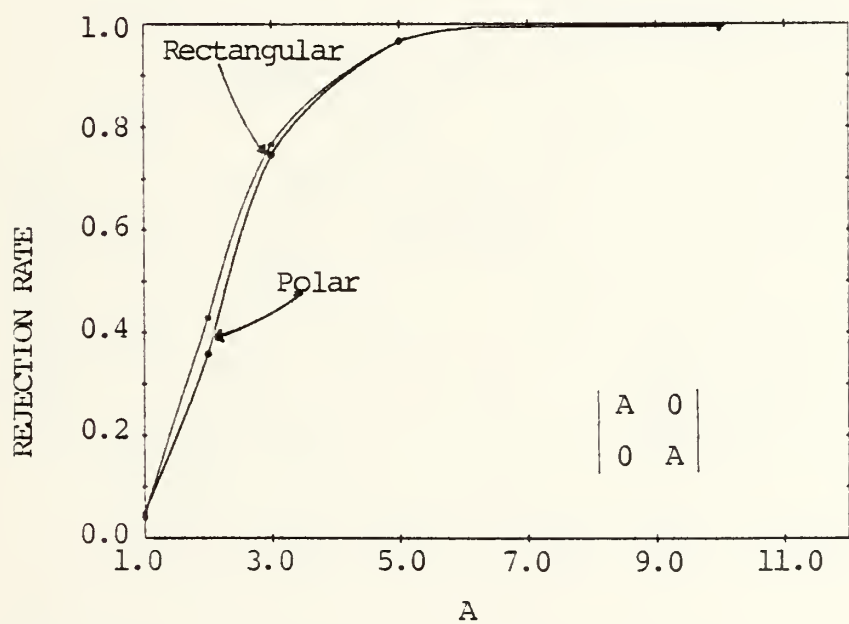
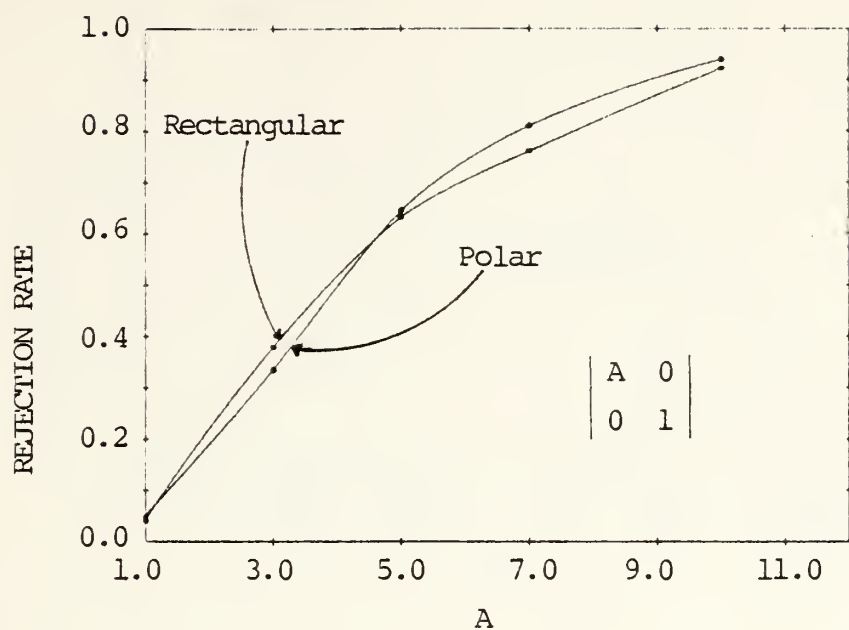


FIGURE 5: POWER CURVES FOR SHIFTS IN VARIANCE (BIVARIATE)

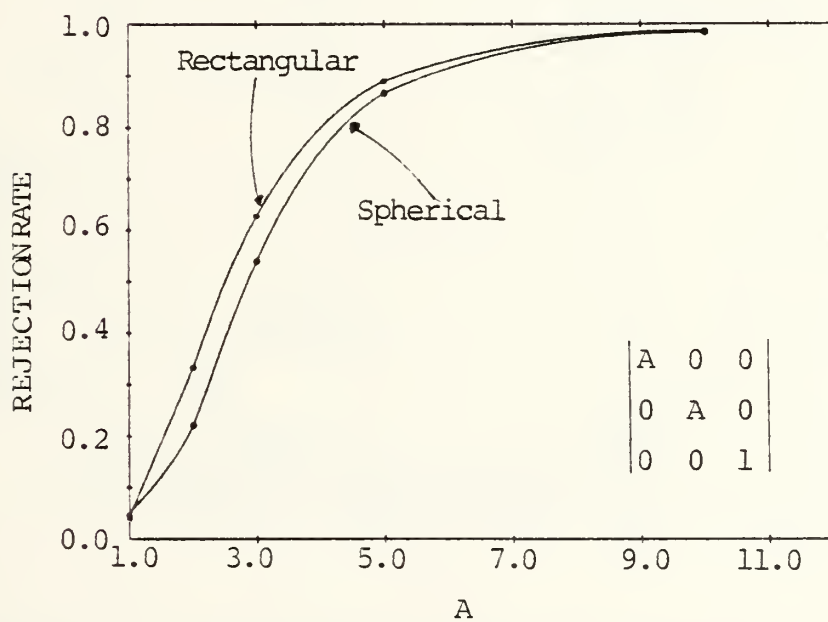
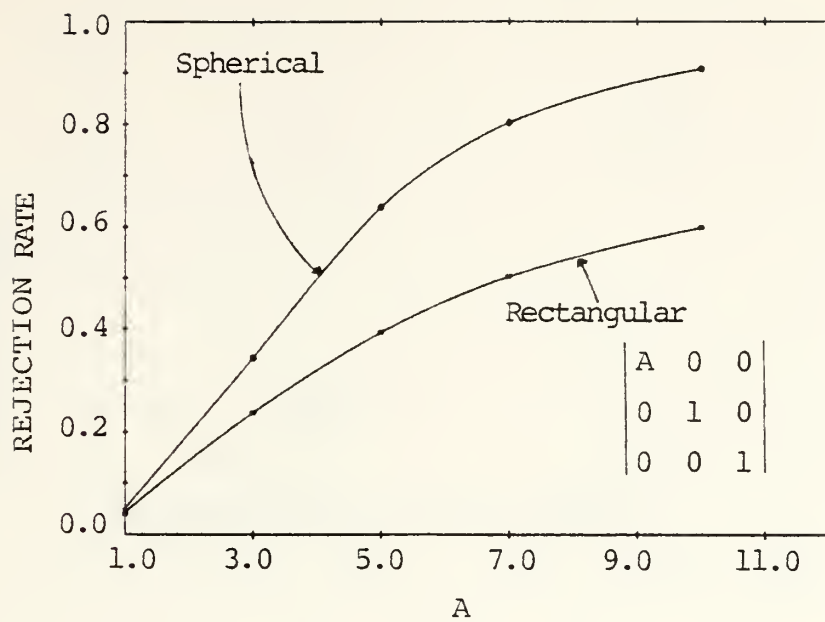


FIGURE 6: POWER CURVES FOR SHIFTS IN VARIANCE (TRIVARIATE)

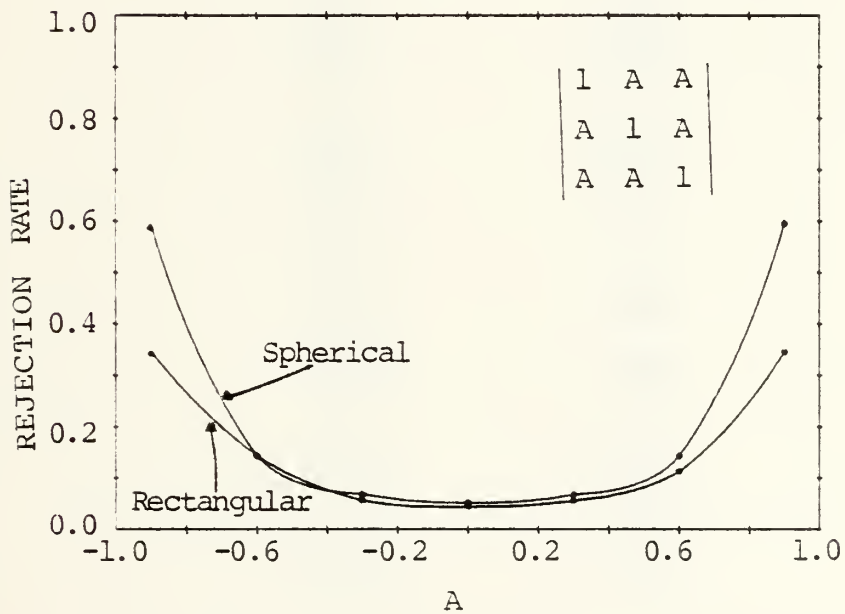
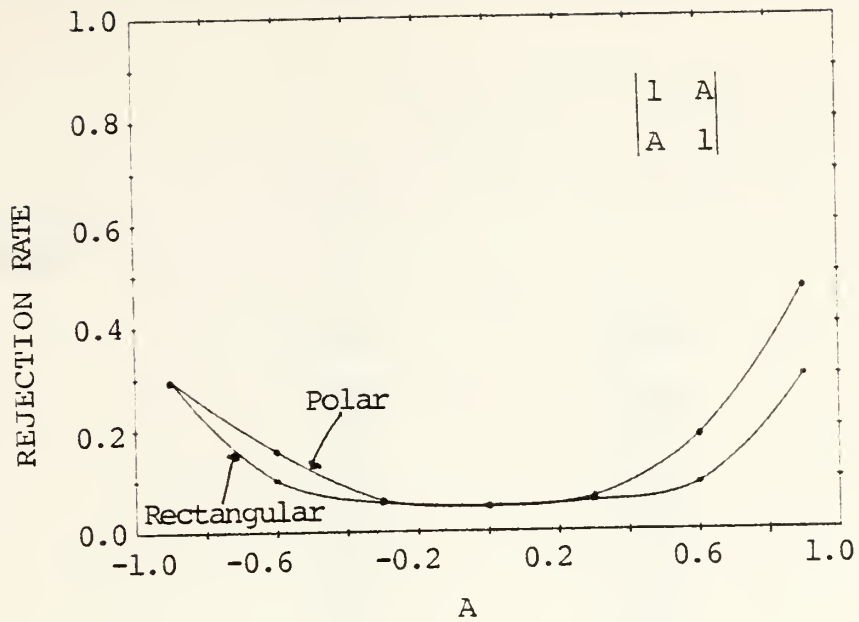


FIGURE 7: POWER CURVES FOR SHIFTS IN COVARIANCE

TABLE VI: REJECTION RATES FOR SHIFTS IN MEAN (BIVARIATE)

N = 20

Critical Value Mean Tested	.01	.05	.10
0	.0094	.0488	.0998
0	.0102	.0482	.0996
-.5	.0566	.1684	.2816
0	.0346	.1096	.2138
.5	.0574	.1710	.2700
0	.0430	.1388	.2298
-.5	.1408	.3164	.4534
-.5	.1038	.2592	.3610
.5	.1294	.3024	.4406
.5	.1230	.3164	.4534
-1	.4357	.6664	.7834
0	.2340	.4484	.6046
1	.4464	.6700	.7842
0	.2444	.6664	.7834
-1	.8382	.9418	.9748
-1	.7780	.9212	.9610
1	.8428	.9418	.9718
1	.6930	.9212	.9610
-2	.9936	.9980	.9996
0	.9926	.9990	.9996
2	.9948	.9998	1.0000
0	.9762	.9950	.9980
-2	1.0000	1.0000	1.0000
-2	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000

BASED ON 5000 REPLICATIONS

FIRST ENTRY--RECTANGULAR

SECOND ENTRY--POLAR

TABLE VII: REJECTION RATES FOR SHIFTS IN MEAN (TRIVARIATE)

N = 20

Critical Value	.01	.05	.10
Mean Tested			
0	.0104	.0440	.0982
0	.0120	.0518	.1048
0			
-.5	.0492	.1502	.2474
0	.0216	.0980	.1632
0			
.5	.0480	.1438	.2484
0	.0280	.1036	.1874
0			
-.5	.1076	.2704	.3990
-.5	.0472	.1516	.2516
0			
.5	.0972	.2424	.3688
.5	.0658	.1912	.3046
0			
-.5	.1826	.3788	.5170
-.5	.0848	.2198	.3584
-.5			
.5	.1738	.3642	.4948
.5	.0848	.2198	.3584
.5			
-1	.3782	.5984	.7212
0	.1184	.2942	.4212
0			
1	.3728	.5984	.7212
0	.1174	.2866	.4234
0			
-1	.7392	.8892	.9410
-1	.3808	.6020	.7338
0			
1	.7400	.8892	.9410
1	.4670	.6918	.7934
0			

TABLE VII (Continued)

Critical Value Mean Tested	.01	.05	.10
-2	.9636	.9872	.9958
0	.7772	.9138	.9916
0			
2	.8992	.9676	.9832
0	.8486	.9448	.9736
0			
-1	.9134	.9778	.9894
-1	.7688	.8744	.9312
-1			
1	.9102	.9746	.9900
1	.7936	.9244	.9598
1			
-2	1.0000	1.0000	1.0000
-2	.9984	.9996	1.0000
0			
2	1.0000	1.0000	1.0000
2	.9998	1.0000	1.0000
0			
-2	1.0000	1.0000	1.0000
-2	1.0000	1.0000	1.0000
-2			
2	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000
2			

BASED ON 5000 REPLICATIONS

FIRST ENTRY--RECTANGULAR

SECOND ENTRY--SPHERICAL

TABLE VIII. REJECTION RATES FOR SHIFTS IN
VARIANCE (BIVARIATE)

N = 20

Critical Values Variance Tested		.01	.05	.10
1	0	.0094	.0488	.0998
0	1	.0102	.0482	.0996
1	0	.1864	.3786	.5150
0	3	.1578	.3342	.4640
2	0	.2228	.4292	.5628
0	2	.1714	.3582	.4928
1	0	.4030	.6322	.7474
0	5	.4286	.6448	.7574
3	0	.5790	.7666	.8580
0	3	.5338	.7450	.8368
1	0	.5640	.7608	.8556
0	7	.6312	.8106	.8856
1	0	.7092	.8618	.9222
0	10	.8088	.9228	.9600
5	0	.8998	.9664	.9832
0	5	.8998	.9665	.9804
10	0	.9956	.9994	.9998
0	10	.9920	.9978	.9988

BASED ON 5000 REPLICATIONS
FIRST ENTRY--RECTANGULAR
SECOND ENTRY--POLAR

TABLE IX. REJECTION RATES FOR SHIFTS IN VARIANCE
(TRIVARIATE)

N = 20

Critical Value Variance Tested			.01	.05	.10
1	0	0	.0104	.0440	.0982
0	1	0	.0120	.0518	.1048
0	0	1			
3	0	0	.0924	.2372	.3626
0	1	0	.1606	.3438	.4736
0	0	1			
2	0	0	.1500	.3330	.4644
0	2	0	.0888	.2208	.3308
0	0	1			
5	0	0	.1940	.3832	.5372
0	1	0	.4146	.6374	.7550
0	0	1			
7	0	0	.2708	.5026	.6332
0	1	0	.6100	.8032	.8792
0	0	1			
2	0	0	.2758	.5012	.6326
0	2	0	.2510	.4538	.5814
0	0	2			
3	0	0	.4140	.6270	.7514
0	3	0	.3240	.5394	.6566
0	0	1			
3	0	0	.6622	.8372	.9038
0	3	0	.6752	.8312	.8966
0	0	3			
10	0	0	.3716	.5980	.7186
0	1	0	.7880	.9078	.9506
0	0	1			
5	0	0	.7558	.8896	.9346
0	5	0	.7256	.8660	.9182
0	0	1			
5	0	0	.9390	.9770	.9872
0	5	0	.9292	.9762	.9852
0	0	5			

TABLE IX (Continued)

N = 20

Critical Value			.01	.05	.10
Variance Tested					
10	0	0	.9572	.9866	.9950
0	10	0	.9470	.9832	.9926
0	0	0			
10	0	0	.9972	.9998	1.0000
0	10	0	.9858	.9970	.9994
0	0	10			

 BASED ON 5000 REPLICATIONS

FIRST ENTRY--RECTANGULAR

SECOND ENTRY--SPHERICAL

TABLE X: REJECTION RATES FOR SHIFTS IN COVARIANCE
(BIVARIATE)

N = 20

Critical Value		.01	.05	.10
Covariance Tested				
1	0	.0094	.0488	.0998
0	1	.0102	.0482	.0996
1	-.3	.0152	.0558	.1068
-.3	1	.0126	.0598	.1274
1	.3	.0126	.0576	.1178
.3	1	.0136	.0656	.1258
1	-.6	.0288	.1008	.1782
-.6	1	.0514	.1576	.2560
1	.6	.0250	.0912	.1702
.6	1	.0648	.1838	.2984
1	-.9	.1166	.2996	.4446
-.9	1	.2378	.2982	.6162
1	.9	.1122	.2996	.4446
.9	1	.2378	.4710	.6042

BASED ON 5000 REPLICATIONS

FIRST ENTRY--RECTANGULAR

SECOND ENTRY--POLAR

TABLE XI: REJECTION RATES FOR SHIFTS IN COVARIANCE
(TRIVARIATE)

N = 20

Critical Value			.01	.05	.10
Covariance Tested					
1	0	0	.0104	.0440	.0982
0	1	0	.0120	.0518	.1048
0	0	1			
1	0	-.3	.0104	.0540	.1076
0	1	0	.0124	.0488	.1066
-.3		1			
1	0	.3	.0106	.0468	.0972
0	1	0	.0124	.0512	.1086
.3	0	1			
1	.3	.3	.0126	.0560	.1112
.3	1	.3	.0162	.0676	.1292
.3	.3	1			
1	0	-.6	.0152	.0740	.1394
0	1	0	.0122	.0584	.1158
-.6	0	1			
1	0	.6	.0148	.0674	.1298
0	1	0	.0128	.0582	.1194
.6	0	1			
1	.6	.6	.0308	.1136	.1960
.6	1	.6	.0432	.1434	.2486
.6	.6	1			
1	0	-.9	.0412	.1358	.2314
0	1	0	.0254	.0974	.1842
-.9	0	1			
1	0	.9	.0402	.1386	.2368
0	1	0	.0258	.1386	.2368
.9	0	1			
1	.9	.9	.1406	.3454	.4942
.9	1	.9	.3646	.5950	.7122
.9	.9	1			

BASED ON 5000 REPLICATIONS

FIRST ENTRY--RECTANGULAR

SECOND ENTRY--SPHERICAL

TABLE XII: REJECTION RATES FOR MULTIPLE SHIFTS
IN MEAN AND VARIANCE-COVARIANCE
(BIVARIATE)

N = 20

Sigma	1 0	1 .6	2 0	2 .849	5 1.34
	0 1	.6 1	0 2	.849 2	1.34 5
Mean					
0	.0488	.0912	.1986	.2500	.9658
0	.0482	.1176	.1522	.2162	.9572
.5	.1710	.2398	.3110	.3702	.9720
0	.1388	.2482	.2498	.3402	.9650
1	.5606	.7346	.6384	.6828	.9820
0	.4348	.5952	.5334	.6316	.9764
1	.9418	.8774	.9350	.8658	.9892
1	.8576	.8588	.8722	.8308	.9840
2	.9998	.9998	.9902	.9950	.9990
0	.9950	.9990	.9772	.9882	.9964

FIRST ENTRY--RECTANGULAR
SECOND ENTRY--POLAR

BASED ON 5000 REPLICATIONS

$\alpha = .05$

TABLE XIII: REJECTION RATES FOR MULTIPLE SHIFTS
IN MEAN AND VARIANCE-COVARIANCE
(TRIVARIATE)

N = 20									
Sigma	1	0	0	1	0	.6	5	0	0
	0	1	0	0	1	0	0	1	0
	0	0	1	.6	0	1	0	0	1
Mean	10	0	.95	5	0	0	10	0	.95
	0	1	0	0	1	0	0	1	0
	0	0	1	0	0	1	.95	0	1
0	.0440			.0674			.5392		
0	.0518			.0582			.4584		
0									
.5	.0480			.1830			.5708		
0	.0280			.1176			.5034		
0									
1	.3728			.6352			.6852		
0	.1174			.2912			.6254		
0									
1	.7400			.9074			.9270		
1	.7392			.7454			.8602		
0									
2	.9982			.9956			.9716		
0	.9726			.9752			.9742		
1									

FIRST ENTRY--RECTANGULAR
SECOND ENTRY--SPHERICAL

BASED ON 5000 REPLICATIONS

$\alpha = .05$

TABLE XIV: REJECTION RATES FOR INCREASING
SAMPLE SIZES (BIVARIATE)

Sample size	20	30	50
Shift			
$\alpha = .01$			
.5	.0574	.0860	.1270
0	.0430	.0564	.0754
.5	.1294	.2026	.3652
.5	.1230	.1418	.2508
1 .3	.0126	.0140	.0176
.3 1	.0136	.0152	.0170
1 0	.1864	.2722	.4522
0 3	.1578	.2244	.3744
.....			
$\alpha = .05$			
.5	.1710	.2170	.2914
0	.1388	.1630	.2238
.5	.3024	.4144	.6030
.5	.3164	.3076	.4826
1 .3	.0576	.0624	.0728
.3 1	.0656	.0624	.0760
1 0	.3786	.4884	.6756
0 3	.3342	.4304	.6016
.....			
$\alpha = .10$			
.5	.2700	.3228	.4256
0	.2298	.2658	.3400
.5	.4406	.5424	.7190
.5	.4534	.4336	.6066
1 .3	.1178	.1174	.1396
.3 1	.1258	.1196	.1422
1 0	.5150	.6132	.7800
0 3	.4640	.5566	.7160

BASED ON 5000 REPLICATIONS

FIRST ENTRY--RECTANGULAR
SECOND ENTRY--POLAR

TABLE XV: REJECTION RATES FOR INCREASING
SAMPLE SIZES (TRIVARIATE)

Sample size Shift	20	30	50
$\alpha = .01$			
.5	.0480	.0680	.1036
0	.0280	.0362	.0526
0			
.5	.1738	.2932	.5040
.5	.0848	.1662	.3428
.5			
1 0 .3	.0106	.0138	.0148
0 1 0	.0124	.0134	.0144
.3 0 1			
3 0 0	.1606	.2054	.3528
0 1 0	.0838	.1138	.2080
0 0 1			
.....			
$\alpha = .05$			
.5	.1438	.1974	.2742
0	.1036	.1256	.1646
0			
.5	.3642	.5118	.7268
.5	.2198	.3588	.5868
.5			
1 0 .3	.0468	.0588	.0656
0 1 0	.0512	.0488	.0540
.3 0 1			
3 0 0	.3438	.3976	.5734
0 1 0	.2074	.2708	.4126
0 0 1			
.....			
$\alpha = .10$			
.5	.2484	.3024	.3876
0	.1874	.2132	.2650
0			
.5	.4948	.6396	.8272
.5	.3584	.4912	.7040
.5			
1 0 .3	.0972	.1142	.1232
0 1 0	.1086	.1030	.1102
.3 0 1			
3 0 0	.4736	.5264	.6880
0 1 0	.3156	.3880	.5450
0 0 1			

BASED ON 5000 REPLICATIONS FIRST ENTRY: RECTANGULAR
SECOND ENTRY: SPHERICAL

APPENDIX A

USER REQUIREMENTS AND INPUT FORMAT FOR PROGRAM FOUTZ

The use of the Computer program contained in Appendix B requires the sample size, number of variates, applicable data and the Multivariate Normal distribution being tested as described by the mean vector and the variance-covariance matrix. The variables containing the required inputs as well as the required input format are as shown below.

DESCRIPTION OF VARIABLES

N-----Sample size
M-----Number of Variables (2 or 3)
SIGMA1-----Variance-Covariance Matrix
 (MxM)
B1-----Mean Vector (Mx1)
X-----Matrix of Sample Data (MxN)

INPUT FORMAT

N,M----- (2I5)
SIGMA1----- (3F12.6) Input M Rows
B1----- (F12.6) Input M Rows
X----- (3F12.6) Input M Rows

Input data is echoed in the output providing a check for correct entry of data as well as is the decomposition of SIGMA1. The Fn statistic as computed by both methods of blocking follows completing the output given for a single run. An example run is given for Trivariate data of sample size 10.

SAMPLE TRIVARIATE RUN

FOUTZ TEST FOR 3 VARIATE NORMAL

THE NUMBER OF OBSERVATIONS = 10

OBSERVATIONS ENTERED AS FOLLOWS

3.170000	7.540000	4.230000
4.160000	5.500000	5.580000
2.330000	2.910000	6.620000
2.530000	3.440000	5.660000
1.990000	2.630000	6.320000
2.260000	2.800000	6.730000
2.630000	0.290000	6.550000
3.440000	4.860000	3.150000
3.500000	4.670000	8.310000
3.580000	3.230000	4.970000

DISTRIBUTION TESTED

COVARIANCE MATRIX

1.000000	1.000000	1.000000
1.000000	3.000000	1.000000
1.000000	1.000000	5.000000

MEAN VECTOR

2.000000
3.000000
4.000000

DECOMPOSITION OF SIGMA

1.000000	0.0	0.0
-0.707107	0.707107	0.0
-0.500000	0.0	0.500000

WITH POLAR OR SPHERICAL COORDINATES

FOUTZ STAT= 0.593289

WITH RECTANGULAR COORDINATES

FOUTZ STAT= 0.556877

APPENDIX B CCOMPUTER PROGRAM FOUTZ

```

C *****
C *          FORTRAN CODE          *
C *          FCUTZ                  *
C *****
C
C N          SAMPLE SIZE
C M          DIMENSION OF EACH VECTOR
C IRAD       N VECTOR DESIGNATING COORDINATE TO CUT ON
C SIGMA1     (M,M) COVARIANCE MATRIX TEST DISTRIBUTION
C B1         (M,1) MEAN VECTOR
C
C          MAIN PROGRAM
C
C PURPOSE:
C          READS IN N,M AND DIMENSIONS
C          LIMITED TO N=50,M=3 AS SET
C
C DIMENSION IRAD(52),VECT(50,6),WKVEC(6),BLOCK(51,12),
$SIGMA1(3,3),B1(3,1),X(50,3),TRAN(3,1),XTT(3,1),C(3,3),
$BLOC(51,12),XTTR(3,1)
C READ(5,990)N,M
990  FORMAT(2I5)
C NN=N+1
C MM=2*M
C NM1=N-1
C DO 10 I=1,N,M
C DO 5 J=1,M
C IRAD(J+I-1)=J
5  CCNTINUE
10 CONTINUE
C CALL DCRIVE(IRAD,VECT,WKVEC,BLOCK,BLOC,SIGMA1,B1,N,M,
$NN,MM,TRAN,XTT,C,XTTR,X)
C STOP
C END
C ..... SUBROUTINE DCRIVE .....
C
C PURPOSE:
C DRIVES PROGRAM AND VARIABLE DIMENSIONS BASED ON
C M AND N. READS IN B1, SIGMA1 AND DATA TO BE
C TESTED. ECHCS INPUT DATA AND PRINTS THE
C RESULTING FN STATISTIC.
C .....
C
C SUBROUTINE DCRIVE(IRAD,VECT,WKVEC,BLOCK,BLOC,SIGMA1,
$B1,N,M,NN,MM,TRAN,XTT,C,XTTR,X)
C DIMENSION IRAC(N),VECT(N,M),WKVEC(3),BLOCK(NN,6),
$SIGMA1(M,M),B1(M,1),TRAN(M,1),X(N,M),BLOC(NN,MM),
$XTT(M,1),C(M,M),XTTR(M,1)
C DO 30 I=1,M
C READ(5,992)(SIGMA1(I,J),J=1,M)
30 CONTINUE
992 FORMAT(3F12.6)
C DO 40 I=1,M

```



```

993 READ(5,993)B1(I,1)
40  FORMAT(F12.6)
CONTINUE
DO 70 I=1,N
READ(5,992)(X(I,J),J=1,M)
70  CONTINUE
CC  ECHO INPUT DATA
WRITE(6,800)M
WRITE(6,801)N
800  FORMAT('1','FCUTZ TEST FOR ',I2,' VARIATE NORMAL')
801  FORMAT('0','THE NUMBER OF OBSERVATIONS =',I3)
804  FORMAT(' ',3F12.6)
805  FORMAT('0','MEAN VECTOR')
806  FORMAT(' ',F12.6)
WRITE(6,791)
791  FORMAT('0','OBSERVATIONS ENTERED AS FOLLOWS')
DO 792 I=1,N
792  WRITE(6,804)(X(I,J),J=1,M)
WRITE(6,807)
WRITE(6,808)
807  FORMAT('0','DISTRIBUTION TESTED')
808  FORMAT('0','COVARIANCE MATRIX')
DO 793 I=1,M
793  WRITE(6,804)(SIGMA(I,J),J=1,M)
WRITE(6,805)
WRITE(6,806)((B1(I,J),J=1,1),I=1,M)
CALL DECOMP(SIGMA,M,C,0)
C  CALL TRANSFORMATION ROUTINES
DO 750 I=1,N
DO 751 J=1,M
751  XTT(J,1)=X(I,J)
752  FORMAT(' ',3X,F12.6)
CALL TRANS(M,XTT,B1,TRAN,C,XTTR)
DO 760 J=1,M
VECT(I,J)=TRAN(J,1)
760  CCONTINUE
750  CONTINUE
CALL BLCKS(VECT,N,NN,M,MM,IRAD,BLOCK,WKVEC)
CALL BLCKR(VECT,N,NN,M,MM,IRAD,BLOC,WKVEC)
C
C BLOCK BY PCLAR OR SPHERICAL COORDINATES
CALL FOUTZ(BLCK,NN,MM,FN,M)
WRITE(6,989)
989  FORMAT('0','WITH PCLAR OR SPHERICAL COORDINATES')
WRITE(6,990)FN
990  FORMAT(' ',F12.6)
C BLOCK BY RECTANGULAR COORDINATES
CALL FOUTR(BLCC,NN,MM,FN,M)
WRITE(6,988)
988  FORMAT('0','WITH RECTANGULAR COORDINATES')
WRITE(6,991)FN
991  FORMAT(' ',F12.6)
RETURN
END

```

```

C .....
C  SUBROUTINE DECOMP
C
C  PURPOSE:
C  DECOMPOSES THE COVARIANCE MATRIX ENTERED.
C  USES CHOLSKY DECOMPOSITION VIA IMSL ROUTINE
C  LUDECP TO PROVIDE A MATRIX C NEEDED BY THE
C  ROUTINE TRANS.
C .....
C
SUBROUTINE DECOMP(SIGMA,M,C,IV)
DIMENSION SIGMA(M,M),C(M,M),A(51),UL(51),L1(6),M1(6)
IJ=1

```



```

DO 100 I=1,M
DO 110 J=1,I
A(IJ)=SIGMA(I,J)
IJ=IJ+1
110 CONTINUE
100 CONTINUE
CALL LUDECP(A,UL,M,D1,D2,IER)
DO 120 I=1,M
DO 130 J=1,M
II=I*(I-1)/2+J
IF(J.LT.I)C(I,J)=UL(II)
IF(J.EQ.I)C(I,J)=1./UL(II)
IF(J.GT.I)C(I,J)=0.0
130 CONTINUE
120 CONTINUE
IF(IV.EQ.1)GO TO 121
CALL INVT(C,M,D,L1,M1)
WRITE(6,799)
799 FORMAT('0','DECOMPCSITION OF SIGMA')
DO 765 I=1,M
765 WRITE(6,700)(C(I,J),J=1,M)
700 FORMAT(' ',3F12.6)
121 CONTINUE
RETURN
END

```

C
C
C
C
C
C
C
C
C

.....

SUBRCUTINE INVT

PURPOSE

INVERT A MATRIX

.....

SUBROUTINE INVT(A,N,D,L,M)

DIMENSION A(1),L(1),M(1)

C
C
C

SEARCH FOR LARGEST ELEMENT

D=1.0

NK=-N

DO 80 K=1,N

NK=NK+N

L(K)=K

M(K)=K

KK=NK+K

BIGA=A(KK)

DO 20 J=K,N

IZ=N*(J-1)

DO 20 I=K,N

IJ=IZ+I

10 IF(ABS(BIGA)-ABS(A(IJ))) 15,20,20

15 BIGA=A(IJ)

L(K)=I

M(K)=J

20 CONTINUE

C
C
C

INTERCHANGE ROWS

J=L(K)

IF(J-K) 35,35,25

25 KI=K-N

DO 30 I=1,N

KI=KI+N

HOLD=-A(KI)

JI=KI-K+J


```

      A(KI)=A(JI)
30  A(JI) =HCLD
C
C      INTERCHANGE COLUMNS
C
35  I=M(K)
    IF(I-K) 45,45,38
38  JP=N*(I-1)
    DO 40 J=1,N
      JK=NK+J
      JI=JP+J
      HOLD=-A(JK)
      A(JK)=A(JI)
40  A(JI) =HCLD
C
C      DIVIDE COLUMN BY MINUS PIVOT (VALUE OF PIVOT
C      ELEMENT IS CONTAINED IN BIGA)
C
45  IF(BIGA) 48,46,48
46  D=0.0
    RETURN
48  DO 55 I=1,N
    IF(I-K) 50,55,50
50  IK=NK+I
    A(IK)=A(IK)/(-BIGA)
55  CONTINUE
C
C      REDUCE MATRIX
C
    DO 65 I=1,N
      IK=NK+I
      HOLD=A(IK)
      IJ=I-N
      DO 65 J=1,N
        IJ=IJ+N
        IF(I-K) 60,65,60
60  IF(J-K) 62,65,62
62  KJ=IJ-I+K
    A(IJ)=HCLD*A(KJ)+A(IJ)
65  CONTINUE
C
C      DIVIDE ROW BY PIVOT
C
    KJ=K-N
    DO 75 J=1,N
      KJ=KJ+N
      IF(J-K) 70,75,70
70  A(KJ)=A(KJ)/BIGA
75  CONTINUE
C
C      PRODUCT OF PIVCTS
C
    D=D*BIGA
C
C      REPLACE PIVOT BY RECIPROCAL
C
    A(KK)=1.0/BIGA
30  CONTINUE
C
C      FINAL ROW AND COLUMN INTERCHANGE
C
    K=N
100 K=(K-1)
    IF(K) 150,150,105
105 I=L(K)
    IF(I-K) 120,120,108
108 JQ=N*(K-1)

```



```

      JR=N*(I-1)
      DO 110 J=1,N
      JK=JQ+J
      HOLD=A(JK)
      JI=JR+J
      A(JK)=-A(JI)
110  A(JI)=HOLD
120  J=M(K)
      IF(J-K) 100,100,125
125  KI=K-N
      DO 130 I=1,N
      KI=KI+N
      HOLD=A(KI)
      JI=KI-K+J
      A(KI)=-A(JI)
130  A(JI)=HOLD
      GO TO 100
150  RETURN
      END

```

..... SUBROUTINE TRANS

PURPOSE: TO TRANSFORM OBSERVATIONS TO $N(0, I)$
 UNDER THE NULL HYPOTHESIS. USES INPUT
 VALUES OF B1 AND THE MATRIX C FROM DECOMP
 TO TRANSFORM THE DATA ENTERED USING,

$$X^* = C(X-B1).$$

```

SUBROUTINE TRANS(M,XTT,B1,TRAN,C,XTTR)
DIMENSION B1(M,1),XTT(M,1),XTTR(M,1),TRAN(M,1),C(M,M)
CALL SUB(XTT,B1,XTTR,M,1)
CALL PRD(C,XTTR,TRAN,M,M,1)
RETURN
END

```

..... SUBROUTINE SUB

PURPOSE
 SUBTRACT ONE MATRIX FROM ANOTHER TO
 FORM RESULTANT MATRIX.

```

SUBROUTINE SUB(A,B,R,N,M)
DIMENSION A(1),B(1),R(1)

```

CALCULATE NUMBER OF ELEMENTS

NM=N*M

SUBTRACT MATRICES

```

DO 10 I=1,NM
10  R(I)=A(I)-B(I)
RETURN
END

```

..... SUBROUTINE PRD

PURPOSE
 MULTIPLY TWO MATRICES TO FORM A
 RESULTANT MATRIX.


```

C .....
C
SUBROUTINE PRC(A,B,R,N,M,L)
DIMENSION A(1),B(1),R(1)
C
IR=0
IK=-M
DO 10 K=1,L
IK=IK+M
DO 10 J=1,N
IR=IR+1
JI=J-N
IB=IK
R(IR)=0
DO 10 I=1,M
JI=JI+N
IB=IB+1
10 R(IR)=R(IR)+A(JI)*B(IB)
RETURN
END
C .....
C SUBROUTINE BLOCKS
C
PURPOSE:
THIS SUBROUTINE TAKES N M-VARIATE VECTORS AND PARTITIONS
A SPACE OF DIMENSION M INTO N+1 STATISTICALLY EQUIVALENT
BLOCKS BY RECCRINO BLOCK COORDINATE RANGES IN A MATRIX
BLOCK BY THE USE OF SPHERICAL OR PLAR COORDINATES AS
CUTTING FUNCTIONS. THE CUTTING COORDINATE USED AT
EACH STEP IS CONTAINED IN A VECTOR IRAD.
C .....
C
SUBROUTINE BLOCKS(VECT,N,NN,M,MM,IRAD,BLOCK,WKVEC)
DIMENSION VECT(N,M),BLOCK(NN,6),IRAD(N),WKVEC(6)
ZL=1.0E-8
BLOCK(1,1)=0.0
BLOCK(1,2)=1000.
BLOCK(1,3)=0.0
BLOCK(1,4)=6.2831853
BLOCK(1,5)=0.0
BLOCK(1,6)=3.1415927
DO 100 J=1,N
TEMP=0.0
DO 110 I=1,M
TEMP=TEMP+VECT(J,I)**2
110 RAD=TEMP**.5
IF(RAD.GT.ZL)GO TO 112
TDEG=6.2831853
PDEG=3.1415927
GO TO 111
112 TARG=VECT(J,1)/((VECT(J,1)**2+VECT(J,2)**2)**.5)
PARG=VECT(J,M)/RAD
IF(TARG.LT.-1.)TARG=-1.
IF(TARG.LT.1.)GO TO 1122
TARG=1.0
1122 IF(PARG.LT.-1.)PARG=-1.
IF(PARG.LT.1.0)GO TO 1123
PARG=1.0
1123 DEG=ACOS(TARG)
IF(VECT(J,2).GT.0.)GO TO 111
IF(DEG.LT.1.5707963)GO TO 113
DEG=DEG+1.5707963
GO TO 111
113 DEG=DEG+4.712389
111 CONTINUE
PDEG=ACOS(PARG)
WKVEC(1)=TEMP

```



```

WKVEC(2)=DEG
WKVEC(3)=PDEG
DO 120 I=1,NN
IF(WKVEC(1).GT.BLOCK(I,2))GO TO 120
IF(WKVEC(1).LT.BLOCK(I,1))GO TO 120
IF(WKVEC(2).LT.BLOCK(I,3))GO TO 120
IF(WKVEC(2).GT.BLOCK(I,4))GO TO 120
IF(M.EQ.2)GO TO 119
IF(WKVEC(3).LT.BLOCK(I,5))GO TO 120
IF(WKVEC(3).GT.BLOCK(I,6))GO TO 120
119 CONTINUE
IBLOCK=I
GO TO 150
120 CCNTINUE
150 JJ=IRAD(J)
BLIM=WKVEC(JJ)
DO 160 I=1,MM
160 BLOCK(J+1,I)=BLOCK(IBLOCK,I)
BLOCK(IBLOCK,2*JJ)=BLIM
BLOCK(J+1,2*JJ-1)=BLIM
100 CONTINUE
RETURN
END

```

```

C ..... SUBROUTINE FCUTZ .....
C
C PURPOSE: TO COMPUTE THE FOUTZ STATISTIC FROM THE
C          BLOCKS DETERMINED BY SUBROUTINE BLOCKS
C          METHCD USES IMSL ROUTINE MDCH TO EVALUATE
C          CHI-SQUARE PROBABILITIES TO EVALUATE THE
C          CLOSED FORM EXPRESSION GIVEN FOR D. THE FN
C          STATISTIC IS GENERATED BY FCUTZ'S CLOSED
C          COMPUTATIONAL FORMULA.
C .....
C

```

```

SUBROUTINE FOUTZ(BLOCK,NN,MM,FN,M)
DIMENSION BLOCK(NN,6),P(51)
DF=FLOAT(M)
TP=0.0
DO 100 I=1,NN
CALL MDCH(BLOCK(I,1),DF,P1,IER)
CALL MDCH(BLOCK(I,2),DF,P2,IER)
P3=P2-P1
P4=(BLOCK(I,4)-BLOCK(I,3))/6.2831853
IF(M.EQ.2)GO TO 85
P5=(COS(BLOCK(I,5))-COS(BLOCK(I,6)))/2.0
GO TO 86
85 P5=1.0
86 P(I)=P3*P4*P5
TP=TP+P(I)
100 CONTINUE
FN=0.0
DO 300 I=1,NN
AMAX=1.0/NN-P(I)
IF (AMAX.LT.0.)GO TO 300
FN=FN+AMAX
300 CONTINUE
RETURN
END

```

```

C ..... SUBROUTINE BLOCKR .....
C
C PURPOSE:
C THIS SUBROUTINE TAKES N M-VARIATE VECTORS AND PARTITIONS
C A SPACE OF DIMENSION M INTO N+1 STATISTICALLY EQUIVALENT
C BLOCKS BY RECORDING BLOCK COORDINATE RANGES IN A MATRIX
C BLOCK BY THE USE OF RECTANGULAR COORDINATES

```



```

C AS CUTTING FUNCTIONS. THE CUTTING COORDINATE USED AT
C EACH STEP IS CONTAINED IN A VECTOR IRAD.
C.....
C
SUBROUTINE BLCKR(VECT,N,NN,M,MM,IRAD,BLOCK,WKVEC)
DIMENSION VECT(N,M),BLOCK(NN,MM),IRAD(N),WKVEC(M)
DO 10 I=1,MM,2
  BLOCK(I,I)=-1000.
  BLCKR(I,I+1)=1000.
10 CONTINUE
DO 100 J=1,N
  DO 110 I=1,M
    WKVEC(I)=VECT(J,I)
110 DO 120 I=1,NN
    DO 130 II=1,M
      IF(WKVEC(II).LT.BLOCK(I,2*II-1))GO TO 120
      IF(WKVEC(II).GT.BLOCK(I,2*II))GO TO 120
130 CONTINUE
      IBLOCK=I
      GO TO 150
120 CONTINUE
150 JJ=IRAD(J)
      BLIM=WKVEC(JJ)
      DO 160 I=1,MM
        BLCKR(J+1,I)=BLOCK(IBLOCK,I)
160 CONTINUE
        BLOCK(IBLOCK,2*JJ)=BLIM
        BLOCK(J+1,2*JJ-1)=BLIM
100 CONTINUE
RETURN
END
C.....
C SUBROUTINE FOUR
C
C PURPOSE:
C TO COMPUTE THE FOURZ STATISTIC FROM THE
C BLOCKS DETERMINED BY SUBROUTINE BLOCKR
C METHOD USES IMSL ROUTINE MDNCR TO EVALUATE
C NORMAL PROBABILITIES TO EVALUATE THE
C CLOSED FORM EXPRESSION GIVEN FOR D. THE FN
C STATISTIC IS GENERATED BY FCUTZ'S CLOSED
C COMPUTATIONAL FORMULA.
C.....
C
SUBROUTINE FOUR(BLOCK,NN,MM,FN)
DIMENSION BLOCK(NN,MM),P(51)
DO 100 I=1,NN
  P(I)=1.0
  DO 200 J=1,MM,2
    CALL MDNCR(BLCKR(I,J),P1)
    CALL MDNCR(BLCKR(I,J+1),P2)
    P3=ABS(P2-P1)
    P(I)=P(I)*P3
    IF(J.NE.MM-1)GO TO 200
200 CONTINUE
100 CONTINUE
  FN=0.0
  DO 300 I=1,NN
    AMAX=1.0/NN-P(I)
    IF(AMAX.LE.0.0)GO TO 300
    FN=FN+AMAX
300 CONTINUE
RETURN
END

```


LIST OF REFERENCES

1. Foutz, R. V., "A Test for Goodness-of-Fit Based on Empirical Probability Measure", Annals of Statistics, v. 8, p. 989-1001, 1980.
2. Naval Postgraduate School Tech Report 53-80-003, A Study of the Properties of a New Goodness-of-Fit Test, by R. Franke and T. Jayachandran, March 1980.
3. Anderson, T. W., "Some Nonparametric Procedures Based on Statistically Equivalent Blocks", Proceedings of the International Symposium on Multivariate Analysis (P. R. Krishnaiah ed.), pp. 5-27, Academic Press, 1964.
4. Op. Cit., 2, pp. 9-11.
5. Naval Postgraduate School Tech Report 55-81-005, The New Naval Postgraduate School Random Number Package LLRANDOMII, by P. A. Lewis and L. Uribe, February 1981.
6. International Mathematical and Statistical Libraries, 7500 Boulevard, Houston Texas, 77036.

INITIAL DISTRIBUTION LIST

	No. Copies
1. Defense Technical Information Center Cameron Station Alexandria, Virginia 22314	2
2. Library, Code 0142 Naval Postgraduate School Monterey, California 93940	2
3. Department of Mathematics Naval Postgraduate School Monterey, California 93940	1
4. Professor C. O. Wilde, Code 53Wm Chairman, Department of Mathematics Naval Postgraduate School Monterey, California 93940	1
5. Professor R. Franke, Code 53Fe Department of Mathematics Naval Postgraduate School Monterey, California 93940	1
6. Professor T. Jayachandran, Code 53Jy Department of Mathematics Naval Postgraduate School Monterey, California 93940	1
7. Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
8. Professor K. T. Marshall, Code 55Mt Chairman, Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
9. Professor H. J. Larson, Code 55La Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
10. Capt Richard J. Linhart Jr. 9 Silverbrook Place Lincroft, New Jersey 07738	2

Thesis

198106

L66345 Linhart

c.1

A study of Foutz's
multivariate good-
ness-of-fit test.

Thesis

198106

L66345 Linhart

c.1

A study of Foutz's
multivariate good-
ness-of-fit test.

thesL66345

A study of Foutz's multivariate goodness



3 2768 002 11803 6
DUDLEY KNOX LIBRARY